AD_____

Award Number: DAMD17-02-1-0549

TITLE: Automated Analysis and Display of Temporal Sequences of Mammograms

PRINCIPAL INVESTIGATOR: Walter F. Good, Ph.D.

CONTRACTING ORGANIZATION: University of Pittsburgh
Pittsburgh, PA 15260

REPORT DATE: July 2006

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 01-07-2006 | Final | 1 JUL 2002 - 30 JUN 2006 |

**4. TITLE AND SUBTITLE**
Automated Analysis and Display of Temporal Sequences of Mammograms

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
DAMD17-02-1-0549

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Walter F. Good, Ph.D.

E-Mail: goodwf@upmc.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

University of Pittsburgh
Pittsburgh, PA 15260

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT:** Although screening for breast cancer has been effective in detecting cancers, it is not clear that the diagnostic information present in sequences of screening exams is currently being utilized. This project has integrated several novel technologies into a system for providing mammographers with information about changing tissue patterns derived from temporal sequences of images. Our hypotheses are: 1) Sequences of screening mammograms contain information about tissue changes that is not otherwise being exploited in the diagnosis of breast cancer; and, 2) Changing tissue patterns can automatically be identified, and correlated with diagnostic questions. To date, all tasks described in the statement of work have been completed. This was a development project and no evaluation of efficacy was proposed, but the subjective assessments of radiologists are being solicited.

**15. SUBJECT TERMS**
screening mammography, mammography, computer-aided diagnosis, image registration and subtraction,electronic mammography display

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | **19b. TELEPHONE NUMBER** *(include area code)* |
| U | U | U | UU | 23 | USAMRMC |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

**Table of Contents**

Title: Automated analysis and Display of Temporal Sequences of Mammograms
Walter F. Good, Ph.D.

## Introduction

Considerable attention has been paid to methods for improving the diagnostic accuracy and efficiency of mammographic screening procedures, including: Assessments of the importance of age of patients to be screened [1]; Appropriate screening intervals [2]; Appropriate screening techniques [3,4]; Recommended number of views per examination [5]; Single vs. double reading [6]; and, Use of prior examinations for comparison [7]. However, in current screening programs mammographers customarily only compare the current exam to one previous exam (e.g., one year prior or two years prior). This is partly for efficiency but also because there is a lack of compelling evidence on how changing tissue patterns relate to the likelihood of developing a malignancy. Retrospective inspection of exams taken in years prior to the detection of a malignancy often indicates the presence of an abnormality. Such anecdotal evidence suggests that it is at least possible that series of screening exams contain additional information that could improve the accuracy of the diagnostic process.

There is also evidence that difference images in certain types of procedures can contain diagnostically useful information, despite the presence of substantial subtraction artifact. In the diagnosis of chest radiography, it has been demonstrated that adding temporally subtracted images to the diagnostic process can, in many situations, significantly improve diagnostic accuracy and reduce mean interpretation time [8]. Difference images obtained by temporal subtraction methods in mammography have properties somewhat similar to images obtained with bilateral subtraction images used in CAD, where it can be shown that differences contain diagnostic information [9-13]. However, like the subtraction of chest images, temporal image subtraction of mammograms is a difficult task due to tissue changes over time, inconsistencies in breast compression, skewing of the three-dimensional breast relative to the image projection during compression and lack of easy landmarks on the breast that can be used to facilitate optimal image registration. Previous attempts to develop acceptable subtraction methods for mammograms, including methods for both bilateral subtraction [9-12] and temporal subtraction [14-16], have had only partial success. As a result, the temporal subtraction of mammograms has not been employed clinically.

In an attempt to improve diagnostic accuracy of screening mammography, this project has adapted and integrated several novel concepts into a system for providing mammographers with information about changing tissue patterns, and the corresponding likelihood of malignancy, derived from temporal sequences of images. Specifically, this includes: 1) A method for locally deforming and registering mammograms [17,18]; 2) Multi-image CAD, which we have previously used to identify corresponding features in ipsilateral views [19]; and, 3) Adjustment of digitized mammograms to account for nonuniform tissue thickness due to breast compression, to enhance soft display of mammograms and to enable more quantitative computerized estimation of tissue composition [20]. Our hypotheses are: 1) Sequences of screening mammograms contain information about tissue changes that have not previously been exploited in the diagnosis of breast

cancer; 2) Changing tissue patterns can automatically be identified, and correlated with diagnostic questions; and, 3) Such information can be made available to mammographers, in an efficient manner, through displays specifically designed for these image sequences.

The main objectives of this project were to provide a system, which can be employed at the discretion of mammographers, to: 1) Normalize images to facilitate comparisons; 2) Apply our previously developed multi-image CAD methods to identify corresponding features between images, and evaluate possible trends for any such features; 3) Detect and classify trends in temporal sequences; 4) Calculate and present various kinds of parameter images; and, 5) Develop a display system for efficiently presenting sequences of exams to mammographers, with the optional ability to emphasize trends by correcting images for differences in exposure, compression, and local misregistration.

## Body

**Overview –** A display system has been created which provides for: 1) Adjusting for the nonuniformity of breast thickness during mammographic exposure; 2) Correcting for differences in exposure between exams; 3) Deforming mammograms so that all images geometrically match the form of the corresponding view in the most recent exam; 4) Measuring local trends in various parameters over time and representing these trends in summary images; 5) Using trend information to attempt to predict risk for malignancy; and, 6) Displaying the normalized mammographic images in a cine mode, along with summary images. The parameters that have been implemented to date are local breast density, local variation of pixel values, and size of localized densities, which provide a general prototype mechanism for other parameters to be added in the future.

Images are first corrected for the nonlinearity of films characteristic curve, and then for their nonuniformity in thickness during breast compression. All images in a temporal sequence are geometrically transformed to match the most recent image. A fully automatic geometric transform was implemented that can map any mammographic image onto a reference image while guaranteeing registration of specific features. Grayscale equivalence is maintained by dividing destination pixel values by the Jacobian of the geometric transformation to correct for local expansion factors. Relative global exposures are equalized with respect to the reference image by applying regression to adjust corresponding pixel value pairs. This transformation guarantees that pectoral lines, nipples and skin lines are all placed in complete registration and that all images appear to acquired at equivalent exposures. After this normalization process, each image in a sequence is processed by feature-detection filters that characterize local tissue properties at each point. At each location, trends over time in parameters are calculated and represented in summary images. For each parameter, the slope of its regression line, and its variance with respect to its regression line, are mapped onto a 2D color scale. An attempt is made to classify trend data as to the likelihood of malignancy, though our current dataset is inadequate for us to demonstrate the efficacy of this. For subjective evaluation by radiologists, the system has been loaded with 240 verified cases for which we have at least three screening exams, however an overall objective evaluation of the system is not within the scope of the current project.

All tasks described in the statement of work have been undertaken and completed successfully. The following paragraphs follow the outline of the statement of work, which we have repeated below for reference.

**Statement of Work**

*Task 1*: Select, verify and digitize cases
- a. A set of 250 sequences of screening mammograms will be selected and verified
- b. All verified cases will be digitized

*Task 2*: Modification of workstation for this project

*Task 3*: Develop display software
- a. Display difference images
- b. Display trends

*Task 4*: Image analysis software development
- a. Correct for characteristic curve
- b. Correct for breast compression
- c. Skin line detection
- d. Breast compression adjustment
- e. Local and global composition
- f. Global composition from 2 views
- g. Local geometric registration
- h. Histogram equalization between images
- i. Modify Multi-View CAD for this project
- j. Create parameter images
- k. Detect and classify trends

*Task 5*: Initial training and testing of classifier

*Task 6*: Interim testing and analysis

*Task 7*: Final training and testing of classifier

*Task 8*: Final processing of all cases

*Task 9*: Data analysis and final report

## Task 1: Select, verify and digitize cases

**1.a. A set of 250 sequences of screening mammograms will be selected and verified –** A total of 250 cases, were collected from the University of Pittsburgh Medical Center or from other local institutions. Of these, 130 are negative and 120 contain malignancies. All cases have been verified, digitized, and assembled into a database. The database has been anonymized by use of an honest broker system, to comply with the requirements of HIPPA.

Case selection was based on a historical prospective approach, in which we examined cases that had at least two prior screening examinations. We only considered cases for which, prior to the latest examination in which a positive abnormality may have been detected, all the mammograms for these patients had been interpreted as negative. Each case consists of two views of a single

breast. Originally, we had expected to select all cases from routine mammographic examinations performed at the UPMC and Magee health care systems, but to find a sufficient quantity of positive cases having the appropriate historical sequence, we have included images from other institutions. Positive cases were drawn from cases that have undergone biopsy. Negative cases, for the most part, are cases that have not been recommended for biopsy, but in some instances, subtle negative cases, which have had biopsies that were negative for the presence of a mass or microcalcification cluster, have been selected.

**1.b. All verified cases will be digitized –** Each selected film has been digitized with a 12 bit modified film digitizer (Lumisys) having an MTF of 27 and 24% at the Nyquist frequency (10 lp/mm) in the X and Y directions, respectively. The digitizer was routinely evaluated and adjusted to maintain spot size, linearity, geometric fidelity, stability and acceptable spatial noise characteristics, and to maintain a linear relationship between pixel value and optical density, as part of our quality assurance procedures.

## Task 2: Modification of workstation for this project

The software and methods discussed below can run on virtually any display system. But for the purposes of this project, an existing 4-monitor display, which was developed in our laboratory, has been adapted. The main consideration in choosing this display was that it could accommodate the minimum resolution requirements of mammography. The individual monitors (Clinton Electronics, Model DS5000) are 21-inch portrait monitors, having a resolution of 2048 x 2560 pixels, 150 ftL light output and P45 phosphor. The displays are driven by a display controller (Dome Imaging Systems, Model Md5/PCX-2) specifically designed for these monitors, and calibrated with a Barten lookup table to accommodate the contrast sensitivity of the human eye. The system has been designed to display four images of a single exam at a time, or a single image and its associated parameter images, or to display sequences of exams sequentially, at a rate determined by viewers. The modifications involved making provisions for the specific display formats (e.g., cine mode, side-by-side with feature alignment) required in this project.

## Task 3: Develop display software

**3.a. Display difference images –** The ability to deform sequentially acquired mammograms so that they can be placed in reasonable registration was a central aspect of this project. In the originally proposed method, which is described in more detail under section **g** of **Task 4** below, we used intersections of nipple axis lines with chest walls as origins of polar coordinate systems for performing mappings between mammograms. We have demonstrated that a modification of this approach is preferable. Specifically, the origin is now defined as the midpoint of the chest wall. The line from this point, passing through the nipple serves as the base for measuring angles. But because this line is not generally perpendicular to the chest wall, a nonlinear monotonic function is used to find corresponding angles between images. Viewers, at their discretion, have the option of either subtracting consecutive images in sequence or subtracting each previous image from a specific (e.g., the most recent) image. Prior to actually performing the subtraction, images are first normalized to correct for tissue thickness, and then geometrically

deformed to match a common template. The details of this procedure, and a comparison of the various techniques, were presented [17,18].

**3.b. Display trends –** Because the parameter images are at a much lower effective resolution that the original mammographic data, these can be displayed side-by-side. The display software has been appropriately modified to allow for alternative display formats. A cine mode has also been implemented on the display system as one means of displaying temporal sequences of image data. This can be used to display either the original mammographic data or the parameter images being generated from each mammographic image. The optimal arrangement of the displayed information depends to a great extent on the relative importance of the various kinds of parameter and trend summary images being generated. We expect a consensus on this to evolve as radiologists gain more experience with the system.

### *Task 4*: Image analysis software development

**4.a. Correct for characteristic curve** – The characteristic curve for each particular film was obtained from its manufacturer. Although these curves are generic, they are sufficiently accurate for this application, given other imperfections in the overall imaging chain. Our routine quality assurance process on the digitizer maintains a linear relationship between pixel value and optical density. The characteristic curve correction algorithm is written to use the inverse of the digitizer characteristic curve to convert pixel values to optical density values, and then us the inverse of the film's characteristic curve to convert optical density to film exposure.

**4.b, 4.d. Correct for breast compression** – The wide dynamic range of mammograms is caused in part by the non-uniformity of breast thickness during breast compression. To correct for this, at each pixel in a mammogram, we determine the relative thickness of the compressed breast as a function of the distance of the pixel from the skin line. Although there are computationally more efficient methods, we measure the distance of each pixel to the skin line with a simple exhaustive search. We divide the range of possible distances into a small number of intervals (typically 32 to 64 depending on image size) and for each interval, all of the tissue pixels whose distance falls within that interval are grouped. For each such group the mean and standard deviation are calculated, and an appropriate function is fitted, with constraints, to the means plus one standard deviation. This "correction" function represents the change in pixel value with respect to distance from the skin line, and actually indicates tissue thickness relative to this distance. For each pixel, the correction function is used to calculate a correction value from the pixel's distance value, and this correction value is used to normalize the pixel value. In the central regions of the breast area, the correction values are 1 and these pixels are left unchanged. Although failure is rather uncommon, it can happen when the skin line detection algorithm fails. Information about breast thickness is retained in a separate file for use in making quantitative measures (e.g., breast density), as discussed below.

Figure 1 shows an example of the result of this process. The two images in the left column are original digitized mammograms, and the two images in the right column are the result of applying the above methods.

**4.c. Skin line detection** – It is necessary to automatically detect the skin line in order to make geometric measurements that are referenced to the skin and nipple. To do this, we first produce a low-resolution version of the image by averaging pixel values over a 16 16 block and then apply the Sobel gradient to the reduced-resolution image. At each point in the image, the original pixel value, the magnitude of the gradient, the direction of the gradient, the radial distance of the pixel from the center of the left edge of the film and a direction value for the surrounding neighborhood are all combined to produce a value which is proportional to the likelihood that the pixel falls on the skin line. These likelihood values are thresholded and then placed in an image where they are processed with morphological operators to eliminate isolated points. The points that remain, which are invariably on the skin line, are mapped back to the high-resolution image and used to fit a Bezier Curve to define a smooth skin line. Although failure is rather uncommon, it can occur for images where the skin line on the film was so dense (i.e., over an optical density of about 4.2) that the digitizer was unable to distinguish between the skin line and the image background. In this event, we did not include the case in this study.

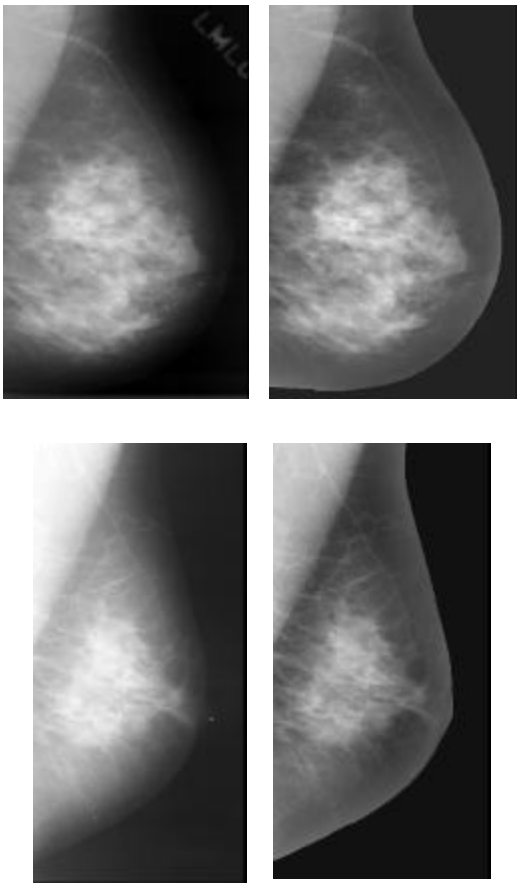

Figure 1: Images on the right are thickness corrected versions of the corresponding images on the left.

**Detection of pectoral muscle** – To detect the pectoral muscle, we employ a template matching with the mammographic image, which is a standard approach, and effective in cases such as this where structures are relatively well defined. The chest wall is not normally visible in the CC view so we assume, in these cases, that it is parallel to, and 0.5 cm beyond the edge of the image.

**4.e. Local and global composition** – After correcting images for characteristic curves of film and then correcting them for nonuniformities due to breast compression, at each point in a view, we estimate % fibroglandular tissue, by deriving features directly from a local histogram and using a neural network, trained with values from breast MRI, to obtain a local estimate of composition. The initial global estimate for a single view is obtained by integrating over the projection of the breast, while accounting for breast thickness at each point in the projection. The local values in each of the two views are scaled relative to the global value averaged from two views, as determined below.

**4.f. Global composition from two views** – As implemented, calculating a global composition for a breast from two views simply involves averaging the values from the CC and MLO views and then adjusting the local values in each of the individual views so that they match the averaged value. Values for individual views are found by the following procedure: 1) Correct

images for characteristic curve of film; 2) Detect skin line; 3) Detect pectoral muscle; 4) Correct for breast compression; 5) Calculate fraction of fibroglandular tissue at each pixel; 6) Multiply this fraction by the breast thickness at the pixel to obtain absolute amount of fibroglandular tissue at pixel; and, 7) Integrate this over the projected area of the breast.

**4.g. Local geometric registration** – To make comparisons between sequentially acquired images more feasible, we adjust the geometry of all images being compared, so that the images can be placed in accurate registration, and we normalize grayscales so that they reflect equivalent exposure conditions. In brief, the subtraction method uses a fully automatic nonlinear transformation that can map any mammographic image onto a template or reference image while assuring concurrent and accurate registration of skin lines, nipples, pectoral muscles and nipple axis lines.

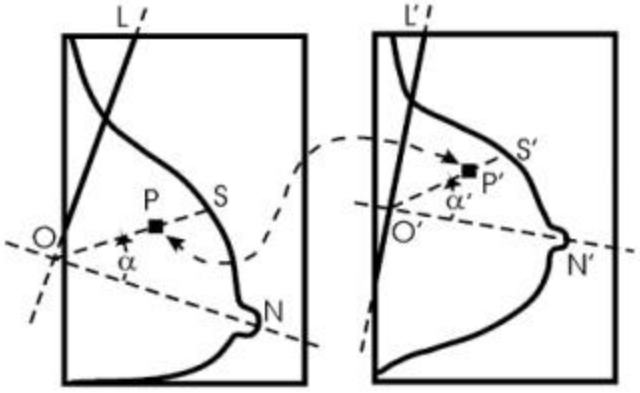


Figure 2. Geometric transformation

The simplest version of the geometric transformation is diagramed in Figure 2. The technique begins by automatically detecting pectoral muscles on MLO views (e.g., *OL* and *O'L'* in Figure 1), skin lines and nipple locations, *N* and *N'*, in Figure 1. Polar coordinate systems are established with the origins, *O* and *O'*, at the intersection of the nipple axes lines (NALs), *ON* and *O'N'*, and lines indicating the pectoral muscles, on MLO views or with lines parallel to, and *0.5 cm* beyond, the edge of the film on CC views. Tissue pixels within a mammogram are identified by a relative polar coordinate, which we define to be the angle, $\alpha$, of their position vector relative to the NAL and their fractional distance between the origin and the skin line. For each pixel *P* in a reference template (e.g., left image in Figure 1), a point *P'*, having the same relative polar coordinate, is found in the image to be deformed. The pixel at location *P* in the template is then given the value of the point at *P'*, as determined by the nearest pixel or through interpolation, after adjusting for the local Jacobian of the transformation at the particular point.

Assuming that the source pixel values (densities) have been corrected for the characteristic curve of the film, so that density and exposure are related by $D = k \ln E$, then the destination pixel value is calculated from the source pixel by

$$p_{dest}\left(T(x), T(y)\right) = p_{source}(x, y) / \left\| J(x, y) \right\| ,$$

where $J(x, y)$ is the Jacobian of the transformation, $\boldsymbol{T}$. This process effectively insures the conservation of linearized density [17-19]. Note that pixels that are reduced in size by the

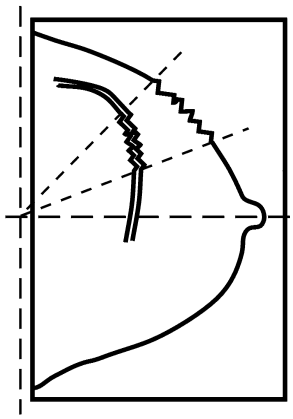

**Figure 3**. Representation of how rough regions of skin line affect transformed features.

geometric transformation (i.e., the Jacobian <1) will have their densities increased and the corresponding film density will be reduced.

Clearly, the geometric transformation maps nipples onto nipples, skin pixels onto skin pixels and chest wall pixels onto chest wall pixels, while interior pixels are deformed correspondingly.

Based on our experience in this project, we developed a second version of this transform, which uses the midpoint of the chest wall as the origin of the polar coordinate system. The line from this point, passing through the nipple serves as the base for measuring angles. But because this line is not generally perpendicular to the chest wall, a nonlinear monotonic function is used to find corresponding angles between images.

The main shortcoming of both of these transformations is that roughness in the skin line description causes a corresponding roughness in interior features in transformed images, as depicted schematically in Figure 3. Our application of using Bezier curves, or other smooth functions, to fit skin line points largely solves this problem.

**4.h. Histogram equalization between images** – Once images have been deformed so that a reasonable correlation between pixels can be assumed, grayscales are corrected by fitting all pixel value pairs for corresponding image points with a linear or nonlinear conversion function (e.g., a 3-segment spline function consisting of two quadratic polynomials smoothly joined by a linear segment). Pixel values of the deformed image are then adjusted by the regression equation, to minimize the sum-of-squares difference, with respect to admissible grayscale changes. A typical distribution of pixel value pairs, with its corresponding correction function, is presented in Figure 4.

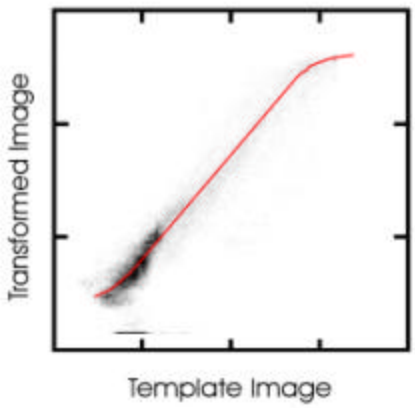


**Figure 4.** Typical correlation between corresponding pixel values in reference image and transformed image after transformation.

Figure 5 shows a typical case where all of these normalization procedures have been applied. The top row shows the original mammograms as digitized, and the middle row depicts the transformed images after the above normalization procedures have been applied.

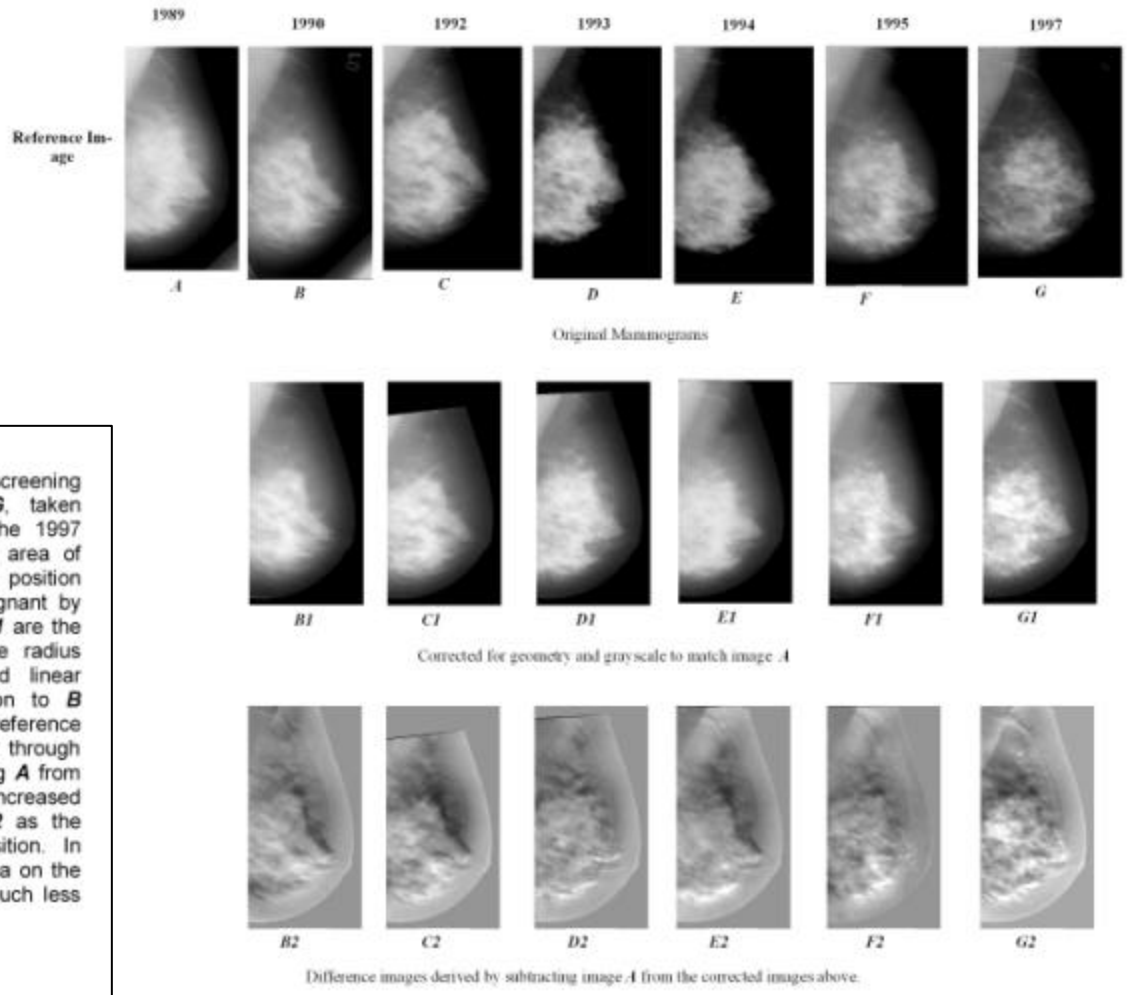


**Figure 5.** Sequence of screening mammograms, ***A*** through ***G***, taken between 1989 and 1997. The 1997 mammogram showed a focal area of shadowing at the 9 o'clock position which was proven to be malignant by biopsy. Images ***B1*** through ***G1*** are the result of applying the relative radius polar coordinate method and linear regression grayscale correction to ***B*** through ***G*** using ***A*** as the reference image. Difference images ***B2*** through ***G2*** are the result of subtracting ***A*** from ***B1*** through ***G1***. The area of increased density is clearly seen in ***G2*** as the bright area in the 9 o'clock position. In contrast, the corresponding area on the original mammogram (***G***) is much less obvious.

**4.i. Modify multi-view CAD for this project** – Our multi-view CAD program, which incorporates methods for comparing features taken at the same time from multiple views, has been modified to permit its use on sequences of images (the same view) taken at different times. Specifically, we included methods for applying linear regression to measure changes in features, with respect to time, across the sequence of images. The main difficulty we have encountered in doing this is that our ability to find corresponding features in a sequence of images depends on our ability to deform the images so that they can be put in geometric registration.

The process begins by using the geometric transformation defined above to deform all images in the sequence to match a single template – usually the first or last image in the sequence. Then pairwise evaluation of all sequential pairs allows corresponding features to be matched and followed throughout the sequence. Parameters characterizing features can also be calculated within each image and followed throughout the temporal sequence.

The modification to our multi-view CAD algorithm involved making provisions for the algorithm to accommodate the geometric correspondence between the geometrically transformed images in the sequence. This is possible in this case because the images represent the same view, though differing in time, but was not possible in our original multi-view CAD where images were of different views.

**4.j. Create parameter images** – A general method for creating parameter images has been implemented. For temporal sequences of images, all images are deformed to the shape of the most recent image. A local feature filter is applied to each image to produce a corresponding image in which pixel values reflect the local magnitude of the particular feature in the original image.

Local features – We have defined a number of local features that could potentially be of value in identifying changing patterns in temporal sequences of mammograms. These include: local breast density, local variation, and various measures of image texture. We also have provided for features related to localized densities detected by multi-view CAD so that these can be followed over time by the same mechanism.

**Local Breast Density** – Local breast density is calculated as described in Task 4.e. above.

**Local Variation** – The two measures of variation that were included were localized fractal dimension and localized standard deviation. In both cases these were calculated within a Gaussian weighted region centered about the given location. Fractal dimension was derived from the slope of the power spectra.

**Characteristics of Localized Densities** – Localized densities detected in image sequences were characterized by size, though other quantitative measures are available from our CAD algorithm, and will likely be considered in the future. When a lesion is detected in any particular image, an attempt is made to follow it forward and backward in the image sequence. For images where a corresponding density is not found, size is set to 0. The measure of size we

employed was the cube root of integrated density (ID), which has previously been developed by us in other contexts [19,21].

The process of calculating ID involves identifying a localized density (e.g., mass) and its boundary and then estimating the average pixel values that would be present if the density did not exist. This estimate of the background density can be obtained by averaging an appropriate area of the image that is outside the region covered by the mass. Pixel values attributed to the mass are defined to be the difference between the actual pixel values in the projected area of the mass and the estimated background value. ID is then calculated as the sum of pixel values attributed to the mass. Theoretically, for fixed magnification and kVp, ID of an object should be invariant with respect to view, breast compression and total exposure under ideal conditions, and we have shown that it is largely invariant under experimental conditions.

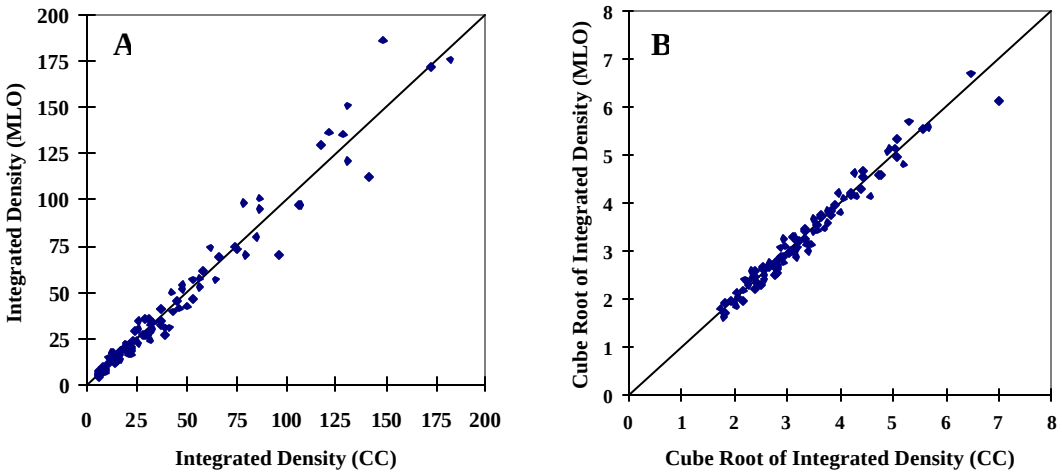


**Figure 6: Scattergrams of (A) integrated density and (B) cube root of integrated density as computed from paired ipsilateral views of 100 masses. The diagonal line is the line of indentity.**

Integrated density is a measure of volume × mean x-ray attenuation. Volume has units of pixels$^3$ or mm$^3$ and attenuation is a unit-less factor, so ID values are expressed in units of volume alone. When mammographers report lesion size, the measure employed (e.g., maximum diameter or effective size) is most often a measure of length. It is misleading to directly compare correlation coefficients between, or errors associated with, measures represented in units of different dimensionality. For the purposes of making such comparisons, we have defined an ID-based measure of effective length as the cube root of integrated density. To verify the stability of this measure, we calculated ID and the cube root of ID for corresponding localized densities appearing in ipsilateral pairs. As depicted in Figure 6, this measure of size shows a high degree of correlation for values calculated for ipsilateral views. In an unpublished study we have also established a high correlation between ID and object volume.

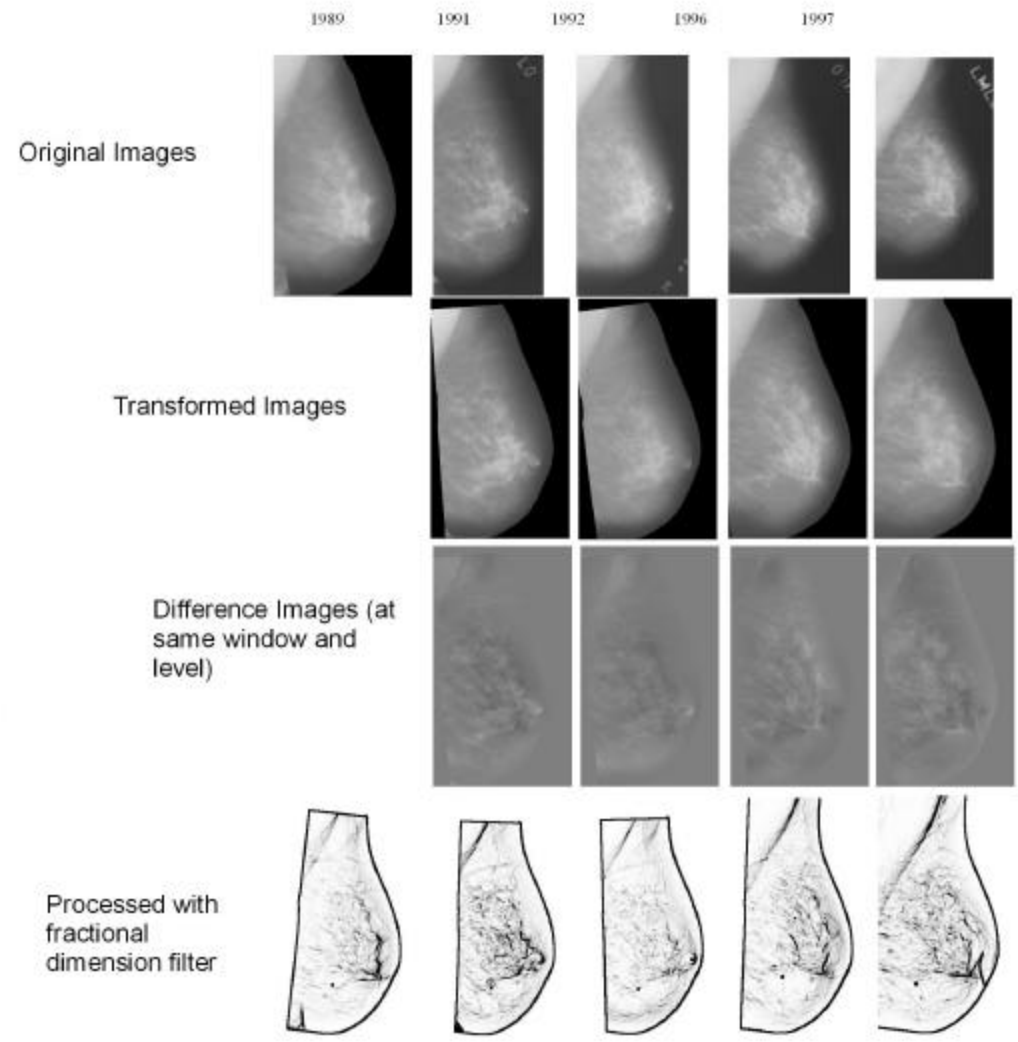


**Figure 7.** Series of five screening exams acquired between 1989 and 1997. Images have been thickness corrected and geometrically transformed using the 1989 image as reference. All images are shown at the same window and level, for quantitative comparison..

**4.k. Detect and classify trends** – For each sequence of parameter files as described above, we have attempted to identify both local and global trends. Direct information about local trends, such as correlation coefficients, is displayable. To create the trend summary image for a feature, at each pixel we calculate a correlation coefficient and the slope of the best-fit line to the sequence of parameter values. In summary images we are currently using hue to represent slope and brightness to represent correlation (with black corresponding to $r^2 = 0$).

In summary images, pixels are assigned values from a scale that ranges from cyan (falling density) to red (increasing density), where brightness indicates the degree of correlation

(uncorrelated sequences of pixel values are shown in black). As a prototype for developing these methods we used local breast density, which was described above. Other filters (e.g., local standard deviation, local fractal dimension) have also been implemented and behave in a similar manner. The results of these methods were presented at the Era of Hope Meeting in 2005. An example of one case that was processed with a fractal dimension filter is shown in Figure 7, and the resulting summary image is shown in Figure 8.

***Task 5*: Initial training and testing of classifier**
***Task 6:* Interim testing and analysis**
***Task 7*: Final training and testing of classifier**

In addition to creating parameter images corresponding to individual features, we have trained a classifier to predict the likelihood of developing a malignancy based on actual features as well as on changes in features. An intense value in one of these images corresponds to a high probability of a malignancy appearing subsequently, at the particular location. Effectively, this was accomplished by adding temporal changes in features, averaged over images, to the feature vector already used in our CAD algorithms. This classification mechanism has been implemented as a neural network, and trained using the training subset of our dataset. This is performed locally for sequences of ipsilateral exams i.e., sequences of pairs of views), and the results of these classifications are displayable. The feature set used as input to the classifier included:

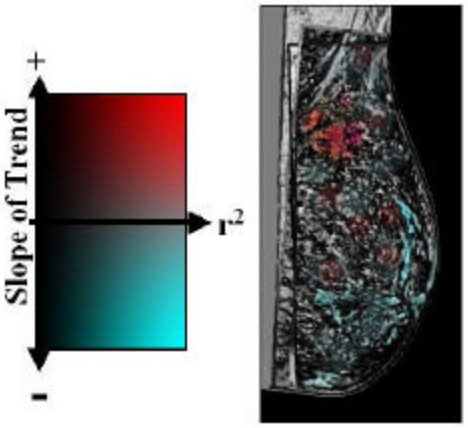


**Figure 8.** Summary image for fractal dimension images in Fig 7. Correlation of corresponding voxels is depicted by brightness and slope of trend by color.

- Local density
- Local trend of changes in density
- Local fractal dimension
- Local trend of changes in fractal dimension
- Probability of malignancy from multi-view CAD

The main significant difference between the final and interim versions of the classification mechanism is that we have now included age as a feature. This was necessary because trends in other parameters, such as breast density, tend to be age dependent. The system has been designed to allow additional features to be easily incorporated, once they are defined and implemented.

We have developed methods and software to measure the contribution of each of the temporal features to the overall performance of the classifier defined above. The neural network classifier is trained and tested on the training set of cases using a jackknife procedure, both with and without a feature of interest. Performance is measured by ROC analysis, with the $A_z$ value being used as a

figure-of merit. The difference in performance, with or without the feature, is taken as a measure of the importance of the particular feature.

The results of the analysis of trends are tentative and are not significant at this point. The main difficulty that arises, in the processes of describing and classifying trends, is that sequences of parameter values do not necessarily follow linear trends. If a nonlinear regression process were used to fit the data then the number of parameters needed to describe trends increases and makes it difficult to devise a classification process that can be applied to the trend data. One possible solution would be to restrict the analysis of trends to only the most recent images in a sequence, but this would neutralize much of the potential diagnostic value of earlier images. An approach that we prefer, but one which was beyond the scope of this project, is to develop a new classification mechanism specifically for analyzing trend data. Most likely this would be a rule based method that can make decisions for the trend of each parameter, and also for correlations between parameters.

Another statistical problem that has arisen in this project is that some of the measures we intended to use are correlated with each other as well as with age, and the strong correlation between malignancy and age make it impossible to find the independent effects of these parameters in a dataset of the size employed in this project. As we continue to acquire images in the future, we expect to overcome this.

Our method for selecting an optimal feature set by using genetic algorithms has been adapted for this project. But because the set of features implemented to date is small the method is not actually useful at this point.

### *Task 8*: Final processing of all cases

All cases in the available dataset were processed by the finalized methods described above and the images have been placed on the workstation developed for this project. While evaluation in terms of efficacy was not described in the original Statement of Work, we are currently soliciting subjective opinions of radiologists.

### *Task 9*: Data analysis and final report

The project was intended to be primarily of a developmental nature, and no quantitative evaluation of overall performance has been performed to date. However, the methods that have been developed were displayed at the Era of Hope Meeting in 2005. Other manuscripts summarizing the work have been prepared and published, or are in the process of being reviewed.

A thorough evaluation of the project, such as by an ROC study, is not feasible at this time because of limitations of our current dataset, and the cost and effort required. However, it will be possible to have the system evaluated subjectively by radiologists within the coming year. Such an evaluation will provide the kind of preliminary results that will make it possible for us to propose a more formal ROC evaluation in the future.

We acquired the image dataset as proposed, though the number of long sequences of screening mammograms, ending with a positive biopsy, was somewhat less than expected. It has turned out that finding biopsy proven positive cases, having long mammographic histories, has been very difficult. This is mainly due to the fact that recommendations for mammographic screening of younger women have historically been ambiguous, and until recently relatively few women routinely have had annual mammographic examinations. The cases we acquired were sufficient for the work proposed in the contract, but this left few positive cases that are available for testing the system. (It is not meaningful to test the system with cases used in its development and optimization.) Discovering generalizable trends by using the methods developed in this project requires these longer exam histories. With this in mind, database acquisition is an ongoing activity in our facility, and will continue with funding from other projects. We are presently trying to collect a sufficient set of these long image sequences for us to perform a more rigorous analysis of the relationship between mammographic trends and risk of malignancy. These additional cases will also increase the power of any evaluation study.

## Key Research Accomplishments

- Selected, verified and digitized 250 cases, and entered them into database.
- Modified workstation for this project.
- Developed display software.
- Developed an innovative registration method for mammograms.
- Developed Image analysis software including software to:
    - Correct for characteristic curve.
    - Correct for breast compression.
    - Detect Skin line.
    - Adjust for breast compression.
    - Calculate local and global composition.
    - Global composition from two views.
    - Perform local geometric registration.
    - Perform histogram equalization between multiple images.
- Identified local and global temporal features.
- Implemented methods for calculating local fractal dimension and integrated density
- Used temporal features to create parameter images.
- Derived trend summary images from the sequences of parameter images.
- Provided a mechanism for optimization of feature sets.
- Processed and installed the dataset of cases on the workstation, for subjective evaluation by radiologists.

## Reportable Outcomes

Manuscripts

Chang YH, Wang XH, Hardesty LA, Chang TS, Poller WR, **Good WF**, Gur D. Computerized assessment of tissue composition from digitized mammograms. Acad Radiol 2002; 9:899-905

Chang YH, Wang XH, Zheng B, **Good WF**, Gur D. Incorporation of negative regions in a knowledge-based computer-aided detection scheme, Proc SPIE 2002; 4684-76

Zheng B, Wang XH, Chang YH, Hardesty LA, Ganott MA, **Good WF**, Gur D. Change of region conspicuity in bilateral mammograms: Potential impact on CAD performance, Proc SPIE 2002; 4684- 87

Wang XH, **Good WF**, Chapman BE, Chang YH, Poller WR, Chang TS, Hardesty LA. Automated assessment of the composition of breast tissue revealed on tissue-thickness-corrected mammography. AJR 2003; 180:257-62.

Zheng B, **Good WF**, Armfield DR, Cohen C, Hertzberg T, Sumkin JH, Gur D. Performance of mammographic CAD schemes optimized using most recent and prior image databases. Acad Radiol 2003; 10:in press

Chang YH, Good WF, Wang XH, Glenn S. Maitz GS, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated Density of a Lesion: A Quantitative, Mammographically-derived, Invariable Measure. Med. Phys. 2003; 30(7):1805-11.

Wang XH, Chapman BE, Good WF. Evaluation of quantitative measures of breast tissue density from mammography with truth from MRI data. Proc SPIE 2003; 5032:82-89.

Good WF, Wang XH, Maitz G. Feature-based differences between mammograms. Proc SPIE 2003; 5032:919-29.

Chang YH, **Good WF**, Wang XH, Glenn S. Maitz GS, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated Density of a Lesion: A Quantitative, Mammographically-derived, Invariable Measure. Med Phys 2003; 30:1805-11.

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Stereo Display for Chest CT. RSNA Dec 2003;

Zheng B, Leader JK, Abrams G, Shindel B, Catullo V, Good WF, Gur D. Computer-aided detection schemes: the effect of limiting the number of cued regions in each case. AJR Am J Roentgenol. 2004 Mar;182(3):579-83.

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Herbert D, Gur D. Stereo Display for Chest CT. SPIE Electronic Imaging 2004; 5291: *In Press*

Zheng B, Gur D, Good WF, Hardesty LA. A method to test the reproducibility and to improve performance of computer-aided detection schemes for digitized mammograms. Med Phys. 2004 Nov;31(11):2964-72.

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Projection Models for Stereo Display of Chest CT. SPIE Medical Imaging 2004; 5367

Good WF, Wang XH, Maitz G. Automated Analysis and Display of Temporal Sequences of Mammograms. Era of Hope 2005; Philadelphia, PA.

Wang XH, Maitz GS, Leader K, Good WF. Real-Time Stereographic Display of Volumetric Datasets in Radiology. SPIE Electronic Imaging 2006, v6055:1A1-1A6.

Databases

We have established a database of verified cases, where each case consists of a sequence of screening mammography exams. While we will continue to accrue cases to this database, it is currently in a form that can be used for analysis.

Proposals applied for

| | |
|---|---|
| US Army MRMC 2003 | Assessing the Interdependence of Tumor Risk, Conspicuity and Breast Tissue Characteristics Derived from Mammography |
| US Army MRMC 2003 | Computerized Analysis and Display of Contralateral Breast Asymmetry |
| US Army MRMC 2003 | Improving Breast Cancer Assessment Through More Accurate Measurement of Mass Size and Growth Rate |
| Komen 2002 | Integrated Density – A Quantitative Measure of Breast Lesion |
| Komen 2003 | Utilization of Mammographic Complexity and Tissue Density in Breast Cancer Detection |
| NIH 2004 | Improving Early Detection of Breast Cancer |
| NIH 2004 | Integrated Density – A Quantitative Measure of Breast Lesion |
| US Army MRMC 2006 | Cine Mode Display for Breast Tomosynthesis |
| PA Dept. of Health | Utilization of Mammographic Complexityfor Improving Risk Assessment and Cancer Detection |

## Conclusions

Progress on this project closely followed the schedule and methods outlined in the original proposal. All tasks described in the statement of work have been undertaken and completed. The display is capable of normalizing and displaying sequences of screening exams, as well as calculating and displaying trend data for various local features.

The main problems we have encountered relate to:

- The insufficient size of our dataset makes it impossible separate the independent effects of multiple variables. This problem will be solved as we continue to enlarge our dataset.

- Our inability to characterize trends in a manner that they can be processed by neural network classifiers. The solution to this involves developing classification mechanisms specifically tailored to trend data.

We remain optimistic about the prospects for the methods developed in this project to have a positive impact on Radiology. The workstation we have developed will make it possible for radiologists to efficiently view a woman's entire mammographic history, as well as view summarized presentations of that history. We expect that the viewing of this additional data will increase radiologists' understanding of the kinds of changes that are occurring in breasts, and the rapidity with which these changes are occurring. Ultimate, this will likely result in increased efficacy for screening mammography.

## References


1. Feig SA, D'Orsi CJ, Hendrick RE. American college of radiology guidelines for breast cancer screening. *AJR* **1998;** 171:29-33.

2. Gordon R. Detection of breast cancer at a small size can reduce the likelihood of metastatic spread: quantitative analysis. *Acad Radiol* **1997;** 4:8-12.

3. Hanchak NA, Kessler HB, MacPerson S. Screening mammography: experience in a health maintenance organization. *Radiology* **1997;** 205:441-445.

4. Moskowitz M. Retrospective reviews of breast screening: what do we really learn from them? *Radiology* **1996;** 199:615-620.

5. Sickles EA. Findings at mammographic screening on only one standard projection: outcomes analysis. *Radiology* **1998;** 209:471-475.

6. Thurfjell EL, Lernevall KA, Taube AS. Benefit of independent double reading in a population-based mammography screening program. *Radiology* **1994;** 191:241-244.

7. Callaway MP, Boggis CR, Astley SA. Influence of previous films on screening mammographic interpretation and detection of breast carcinoma. *Clin Radiol* **1997;** 52:527-529.

8. Difazio MC, MacMahon H, Xu XW, Tsai P, Shiraishi J, Armato SG, Doi K. Digital chest radiography: effect of temporal subtraction images on detection accuracy. *Radiology* **1997**; 202:447-452.

9. FF Yin, ML Giger, K Doi, CE Metz, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images. *Med Phys* **1991**; 18:955-963.

10. FF Yin, ML Giger, CJ Vyborny, K Doi, RA Schmidt. Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. *Invest Radiol* **1993**; 6:473-481.

11. FF Yin, ML Giger, K Doi, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. *Med Phys* **1994**; 21:445-452.

12. B Zheng, YH Chang, D Gur. Computerized detection of masses from digitized mammograms: Comparison of single-image segmentation and bilateral image subtraction. Acad Radiol **1995**; 2:1056-106

13. Mendez AJ, Tahocas PG, Loda MJ. Computer-aided diagnosis: automatic detection of malignant masses in digital mammograms. *Med Phys* **1998;** 25:957-964.

14. M Sallam, K Bowyer. Registering time sequences of mammograms using a two-dimensional image unwarping technique. SIWDM **1994**.

15. D Brzakovic, N Vujovic, M Neskovic, P Brzakovic, K Fogarty. Mammogram analysis by comparison with previous screenings. SIWDM **1994**.

16. S. Sanjay-Gopal, HP Chan, N Petrick, TE Wilson, B Sahiner, MA Helvie, MM Goodsitt. A regional mammogram registration technique for automated analysis of interval changes of breast lesions. Proc. SPIE 3338:118-123.

17. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Generalized procrustean image deformation for subtraction of mammograms. *Proc SPIE on Medical Imaging* **1999;** 3666-167.

18. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Image modification for displaying temporal sequences of mammograms. *Proc SPIE on Medical Imaging* **2000;** 3976 (in press).

19. Good WF, Zheng B, Chang YH, Wang XH, Maitz GS, Gur D. Multi-image CAD employing features derived from ipsilateral mammographic views. *Proc SPIE on Medical Imaging* **1999;** 3661-47.

20. Wang XH, Zheng B, Chang H, Good WF. Correction of digitized mammograms to enhance soft display amd tissue composition measurement. *Proc SPIE on Medical Imaging* **2001**; Vol 4319.

21. Chang YH, Good WF, Wang XH, Glenn S. Maitz GS, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated Density of a Lesion: A Quantitative, Mammographically-derived, Invariable Measure. Med. Phys. 2003; 30(7):1805-11.

## List of Personnel

Waltr F. Good, Ph.D.

Xiao Hui Wang, M.D., Ph.D.

Brian Chapman, Ph.D.

Yuan-Hsiang Chang, Ph.D.

**Keywords:** Mammography; Screening Mammography; Computer-Aided Diagnosis; Image Registration and Subtraction; Electronic Mammography Display